



Predicting facial characteristics from complex polygenic variations

Fagertun, Jens; Wolffhechel, Karin Marie Brandt; Pers, Tune ; Nielsen, Henrik Bjørn; Gudbjartsson, Daniel ; Stefansson, Hreinn; Stefansson, Kari; Paulsen, Rasmus Reinhold; Jarmer, Hanne Østergaard

Published in:
Forensic Science International: Genetics

Link to article, DOI:
[10.1016/j.fsigen.2015.08.004](https://doi.org/10.1016/j.fsigen.2015.08.004)

Publication date:
2015

Document Version
Peer reviewed version

[Link back to DTU Orbit](#)

Citation (APA):
Fagertun, J., Wolffhechel, K. M. B., Pers, T., Nielsen, H. B., Gudbjartsson, D., Stefansson, H., Stefansson, K., Paulsen, R. R., & Jarmer, H. Ø. (2015). Predicting facial characteristics from complex polygenic variations. *Forensic Science International: Genetics*, 19, 263-268. <https://doi.org/10.1016/j.fsigen.2015.08.004>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Accepted Manuscript

Title: Predicting facial characteristics from complex polygenic variations

Author: Jens Fagertun Karin Wolffhechel Tune Pers Henrik
B. Nielsen Daniel Gudbjartsson Hreinn Stefansson Kári
Stefansson Rasmus R. Paulsen Hanne Jarmer



PII: S1872-4973(15)30059-4
DOI: <http://dx.doi.org/doi:10.1016/j.fsigen.2015.08.004>
Reference: FSIGEN 1403

To appear in: *Forensic Science International: Genetics*

Received date: 9-4-2015
Revised date: 15-7-2015
Accepted date: 13-8-2015

Please cite this article as: Jens Fagertun, Karin Wolffhechel, Tune Pers, Henrik B. Nielsen, Daniel Gudbjartsson, Hreinn Stefansson, Kári Stefansson, Rasmus R. Paulsen, Hanne Jarmer, Predicting facial characteristics from complex polygenic variations, *Forensic Science International: Genetics* (2015), <http://dx.doi.org/10.1016/j.fsigen.2015.08.004>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Highlights

- We predict facial traits from genetic variants of 1,300 individuals
- We use information from large sets of SNPs
- Six facial traits, facial width among others, are predicted significantly
- This work is a promising addition to previous research into facial trait prediction

Predicting facial characteristics from complex polygenic variations

Jens Fagertun^{*a}, Karin Wolffhechel^{*b}, Tune Pers^{b,d,e}, Henrik B. Nielsen^b,
Daniel Gudbjartsson^c, Hreinn Stefansson^c, Kári Stefansson^c, Rasmus R.
Paulsen^a, Hanne Jarmer^b

a. Department of Applied Mathematics and Computer Science, Technical University of Denmark, Lyngby, Denmark

b. Center for Biological Sequence Analysis, Department of Systems Biology, Technical University of Denmark, Lyngby, Denmark

c. deCODE genetics, Reykjavik, Iceland

d. Broad Institute of MIT and Harvard, Cambridge, USA

e. Division of Endocrinology and Center for Basic and Translational Obesity Research, Boston Children's Hospital, Boston, USA

* These authors contributed equally to this work

The corresponding author is Karin Wolffhechel, karinw@cbs.dtu.dk

1 Abstract

2 Research into the importance of the human genome in the context of facial appearance is receiving
 3 increasing attention and has led to the detection of several Single Nucleotide Polymorphisms (SNPs) of
 4 importance. In this work we attempt a holistic approach predicting facial characteristics from genetic
 5 principal components across a population of 1,266 individuals. For this we perform a genome-wide
 6 association analysis to select a large number of SNPs linked to specific facial traits, recode these to genetic
 7 principal components and then use these principal components as predictors for facial traits in a linear
 8 regression. We show in this proof-of-concept study for facial trait prediction from genome-wide SNP
 9 data that some facial characteristics can be modeled by genetic information: facial width, eyebrow width,
 10 distance between eyes, and features involving mouth shape are predicted with statistical significance
 11 ($p < 0.03$).

12 2 Keywords

13 facial trait prediction; visible trait prediction; normal trait variation; evolutionary genetics; genetic
 14 association.

15 3 Introduction

16 Facial features are one of our most distinguishing visible traits, and known to be modeled by combined
 17 genetic, epigenetic and environmental factors [1]. An example of this are monozygotic twins, often
 18 so similar in appearance that they can be difficult to tell apart, which concludes that at least partial
 19 reconstruction of facial features from DNA should be possible.

20 Genotypic profiling based on Single Nucleotide Polymorphisms (SNPs) has been successfully implemented
 21 to predict a person's amount of freckling, presence of moles, hair texture and skin color [2–8]. In forensic
 22 science genotyping has been used to predict an individual's eye color, hair color, sex and ancestry with
 23 high accuracy [9, 10]. Prediction of facial features from DNA would be very useful in forensic science,
 24 since this would make it possible to narrow down a list of suspects for a particular crime based on facial
 25 appearance alone.

26 However, this aim is difficult to implement, since Genome-Wide Association Studies (GWAS) have

revealed many traits as inherently polygenic [11–13]. One example of this is human height, now believed to be shaped by thousands of SNPs [14]. The search for decoding the genetic modeling of facial traits has at this time only taken the first steps: three SNPs affecting nose width, nasion position and face width have been detected [15, 16]; Liu *et al.* identified five SNPs associated with specific facial phenotypes extracted from three dimensional (3D) magnetic resonance images and two dimensional (2D) portrait images [17]; and a study conducted by Claes *et al.* established 24 of 76 candidate SNPs in known craniofacial genes as having significant effects on facial variation [18]. Nonetheless, lack of SNPs with strong effects and a low degree of explained facial variance reveal facial phenotypes as polygenic. Therefore, other methods are needed to increase the chance of success for finding a translation from genetics to facial features.

In this work we attempt to predict 2D facial characteristics from genetic variants across a population of 1,266 individuals to establish the presence of a signal in our DNA for the coding of facial features. This is a proof-of-concept study focusing on genome-wide SNP data instead of individual SNPs for the prediction of facial traits. We construct a statistical shape model of the face by Principal Component Analysis (PCA). The facial shape principal components are used as phenotypes in a GWAS to select for associated SNPs, whereafter a PCA is performed on these selected SNPs to construct corresponding statistical models for the genetic variation. The genetic principal components are then used to predict a given facial shape component by linear regression in a repeated stratified nested cross-validation design.

4 Materials and Methods

4.1 Ethics Statement

Subjects were recruited from a large genotyped sample [19]. Encrypted identifiers of genotyped subjects were decrypted by a representative of the Icelandic Data Protection Authority and subjects were recruited to the study by a clinic overseen by the Icelandic Data Protection Authority. Psychologists and nurses phenotyping the participants were blind to genotype. Those working with the genetic data were blind to personal identifiers and could only work on the encrypted data set. Only a representative of the Data Protection Authority of Iceland holds the key for encrypting and decrypting the personal identifiers. Genotypes are only linked to encrypted identifiers. Approval for this study was obtained from the National Bioethics Committee of Iceland (VSNb20090900004) and the Icelandic Data Protection Authority. Written informed consent was obtained from all participants before blood samples or phenotypic data were obtained.

55 All sample identifiers were encrypted in accordance with the regulations of the Icelandic Data Protection
56 Authority.

57 4.2 Overview

58 This study focuses on the prediction of facial features from genome-wide SNP data. Over twenty million
59 genotyped and imputed SNPs were available and therefore feature selection by genome-wide association
60 was implemented to ensure genetic models being built with SNPs associated with a given shape component.
61 An overview of our method is presented in Figure 1.

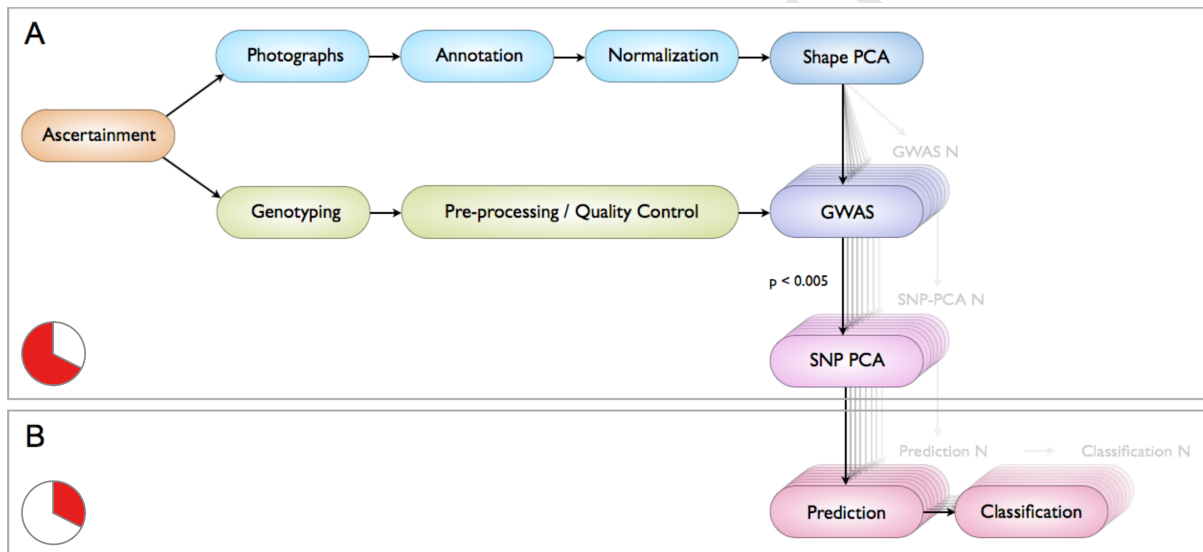


Figure 1. Flow chart giving an overview of our method. A: Participant photographs are annotated, normalized and a PCA is performed on the face shapes resulting in 37 shape components (95 % explained variance). Participant genotypes are pre-processed and a GWAS is run against each shape component ($N=37$) on two-thirds of the data. SNPs with an association p -value < 0.005 are selected and a PCA is performed on these, resulting in about 670 genetic components per shape component. B: The remaining one-third of data is used to train and assess the model by repeated stratified nested 40-fold cross-validation. The final model is assessed by the multiple correlation coefficient on the test set.

62 4.3 Ascertainment

63 The ascertainment of Icelandic participants has been described in detail elsewhere [19]. Of these 1,338 were
64 acquired with facial image data, 72 participants were removed due to non-Icelandic ethnicity, non-neutral
65 facial expression, or non-frontal face direction to the camera, resulting in a remainder of 1,266 participants.

66 All participants were between 18 and 72 years of age (57% women).

67 **4.4 Genotyping**

68 Participants were genotyped using IlluminaHumanHap and IlluminaOmniExpress arrays and long range
69 phased for efficient imputation of markers as described earlier [20]. To shortly recap, SNPs identified
70 and genotyped through sequencing were imputed into all Icelanders who had been phased with long
71 range phasing using the same model as used by IMPUTE [21]. SNPs were excluded if they (i) had
72 yield less than 95%, (ii) had MAF less than 1% in the population or (iii) had significant deviation from
73 Hardy-Weinberg equilibrium ($p < 0.001$). All samples with a call rate below 98% were excluded from the
74 analysis. Genotyping was performed at deCODE genetics in Reykjavik, Iceland.

75 **4.5 Image acquisition**

76 Frontal facial images were recorded in a controlled environment at enCODE and Landspítali University
77 Hospital in Reykjavik, Iceland. At both sites identical photo studios were set up and all images were
78 taken with a Canon IXUS 95, 10 megapixels camera, mounted on a custom-built rig. All participants were
79 recorded sitting on a height-adjustable chair. Lighting was not controlled and therefore we subsequently
80 chose to exclude texture information. Images were corrected for distortions, mainly introduced by the
81 camera lens, by use of camera calibration [22].

82 **4.6 Annotation of facial landmarks**

83 Image analyses were performed in Matlab 2010b [23] and facial images were annotated by in-house
84 developed image analysis software employing a semi-automatic annotation scheme. First, face and eyes
85 were automatically detected using the Viola-Jones object detection algorithm [24] and then an Active
86 Appearance Model placed the remaining annotation points [25]. The annotation scheme consisted of 73
87 anatomical landmarks and pseudo-landmarks and was similar to the one used in [26], an example is shown
88 in Figure 2. A trained operator manually revised and adjusted all annotations.

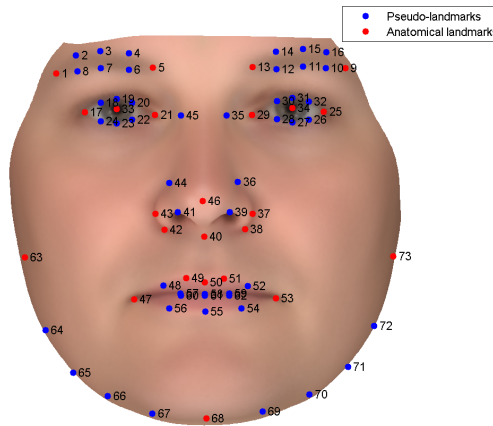


Figure 2. 73 points annotation scheme. An example of the set of annotation points to determine the shape model. Anatomical landmarks are points determining homologous parts in a face; like the tip of the nose. Pseudo-landmarks are points located between anatomical points that are needed to construct a precise shape model.

4.7 Generating shape components

A Procrustes Analysis was performed on the annotated faces to align and scale them due to missing absolute distances in the 2D images [27]. In a procrustes analysis optimal transformation of annotation points is found to minimize distance between points of the same class; e.g. the tip of the nose (Figure 3). The mean shape was subtracted from these aligned and scaled shapes and a PCA was performed, resulting in a statistical facial shape model with principal components restricted to account for 95% of the variance in the original data. Facial shape principal components are further-on referred to as shape components.

4.8 Generating genetic components

A GWAS was performed for each shape component, adjusted for gender and age effects, on two-thirds of the data, 827 individuals, to detect associated SNPs. To account for relatedness and stratification within samples the method of genomic control based on chip markers was applied. SNPs with association p-values < 0.005 were selected and genotypes for the selected SNPs were coded to an additive model with missing values coded as heterozygous. A PCA was applied to construct statistical genetic models, again with principal components restricted to account for 95% of the variance in the original data, and further-on referred to as genetic components. Data from the 827 individuals used for the model generation was discarded and the remaining one-third of data, 439 individuals, was used for the prediction. Genotypes for

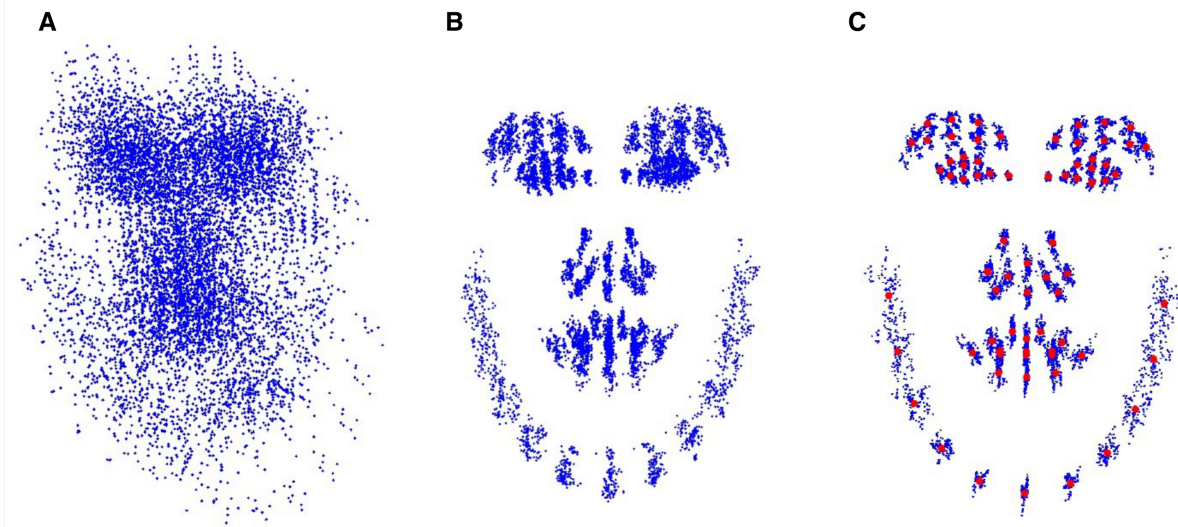


Figure 3. Procrustes analysis. A: The original annotation points/landmarks. B: Translation of the center of gravity into origin. C: Result of the Procrustes analysis; the optimal rigid transformations (rotation, scaling and translation) that minimize the distance between points of the same class (e.g. left pupil). The mean shape from the Procrustes analysis is shown in red.

the selected SNPs for the 439 individuals were projected into the PCA spaces of the previously generated genetic models.

4.9 Prediction of facial traits

Shape components were predicted from genetic components by training and validating linear regression models by repeated stratified nested 40-fold cross-validation, where the outer loop estimates performance and the inner loop selects the best number of genetic components for the prediction. Due to a setup with much fewer observations than genetic components feature selection in the inner loop was performed by first ranking all genetic components based on their correlation to their corresponding shape component in the training set [28]. Up to fifty of the highest ranked genetic components were used as predictors and the best number of predictors was chosen in the inner loop. Linear regression models were trained in the outer loop and performance was evaluated on the test set by the correlation coefficient, r , between predicted and target values. Cross-validation was repeated a hundred times to report results with standard deviations. The prediction was additionally done for each gender separately. Permutation tests were performed with 10,000 rounds, resampling shape component values each round, to establish the null distribution of the test statistic and thereby assess the statistical significance of the predictions. The analyses in this section

were performed in R v. 3.1.1 [29].

5 Results

In this study it was our goal to predict holistic facial characteristics from a number of complex genetic components.

For this, participants' images were annotated, procrustes aligned and scaled, whereafter facial shape components were extracted by PCA. This resulted in 37 shape components, each describing holistic aspects of a face (Figure S1). We ran a GWAS for each shape component controlling for age and gender on two-thirds of the data, 827 individuals, to select for associated SNPs. As expected, we did not find SNPs with genome-wide significant p-values ($< 10^{-8}$) due to our small sample size and due to the polygenic nature of facial traits where many SNPs have minor effects. We chose a threshold of $p < 0.005$ to capture a large number of possibly associated SNPs across the genome; other thresholds were not tested. Thirty-seven genetic models, one for each shape component, were built by a PCA from the selected SNPs ($p < 0.005$, $\sim 95,000$ SNPs). PCA is often done on genetic data from different populations to control for population stratification by use of the first few principal components. Our data is from one population only and therefore a PCA is in this case appropriate for dimension reduction to capture differences in individuals' genetic data accounting for differences in phenotypes. We retained genetic PCs explaining cumulatively up to 95% of the variance (Figure S2). After this initial model generation the above two-thirds of data were discarded to ensure full separation between feature selection and prediction. Next, genotypes for the associated SNPs of the remaining 439 individuals were projected into the PCA spaces of the 37 generated genetic models; one for each shape component. The genetic components were then used to predict shape components in a linear regression model by repeated stratified nested 40-fold cross-validation, where the inner cross-validation selects for the best number of predictors.

Prediction was done on three groups: all 439 individuals, only men and only women. This approach was chosen due to different facial compositions between genders; e.g. it has been shown that men have a larger face width at the cheeks and thicker eyebrows [30]. Adding gender as a covariate would therefore introduce an upwards bias on performance for at least some of the shape components. For each shape component and each group we ran permutation tests with 10,000 rounds to establish the null distribution, giving us a measure of statistical significance for our predictions. We found shape components 1, 3,

Table 1. Shape components for each group predicted with statistical significance.

All, n=439			
Shape component	$r \pm \sigma$	# genetic components	p-value
3	0.21 ± 0.02	6	9.0×10^{-3}
Men, n=231			
Shape component	$r \pm \sigma$	# genetic components	p-value
27	0.28 ± 0.03	2	1.1×10^{-2}
1	0.21 ± 0.03	4	4.4×10^{-2}
Women, n=208			
Shape component	$r \pm \sigma$	# genetic components	p-value
11	0.31 ± 0.00	1	8.0×10^{-3}
16	0.27 ± 0.03	3	1.8×10^{-2}
28	0.25 ± 0.04	1	2.4×10^{-2}

Shape components predicted with statistical significance based on permutation tests. The multiple correlation coefficient, r , is reported with standard deviations together with the number of genetic components used in the prediction and the p-value gathered from the permutation test.

11, 16, 27 and 28 to be predicted with statistical significance (Table 1). However, correlations between observed and predicted shape component scores did not exceed values of 0.31, revealing predictions as only approximate.

In Figure 4 we visualize the changes explained by the shape components. Shape component 3 is seen to mainly explain face width. For women component 11 mainly describes fullness of lips, component 16 distance between eyes and eyebrow width and component 28 eye size. For men component 27 describes the shape of the mouth, while component 1 explains face width.

6 Discussion

In this work we demonstrate that facial characteristics can be predicted to some degree solely from genetic information. This is a proof-of-concept study with the focus on genome-wide SNP data instead of on individual SNPs. Despite a modest number of individuals we found six holistic facial characteristics predicted with statistical significance. We find face width (found previously to be associated to specific SNPs by [15,17]), fullness of female lips, and a slight variation in mouth width in men as the features with the greatest potential in the context of facial trait prediction. For women we additionally find eye

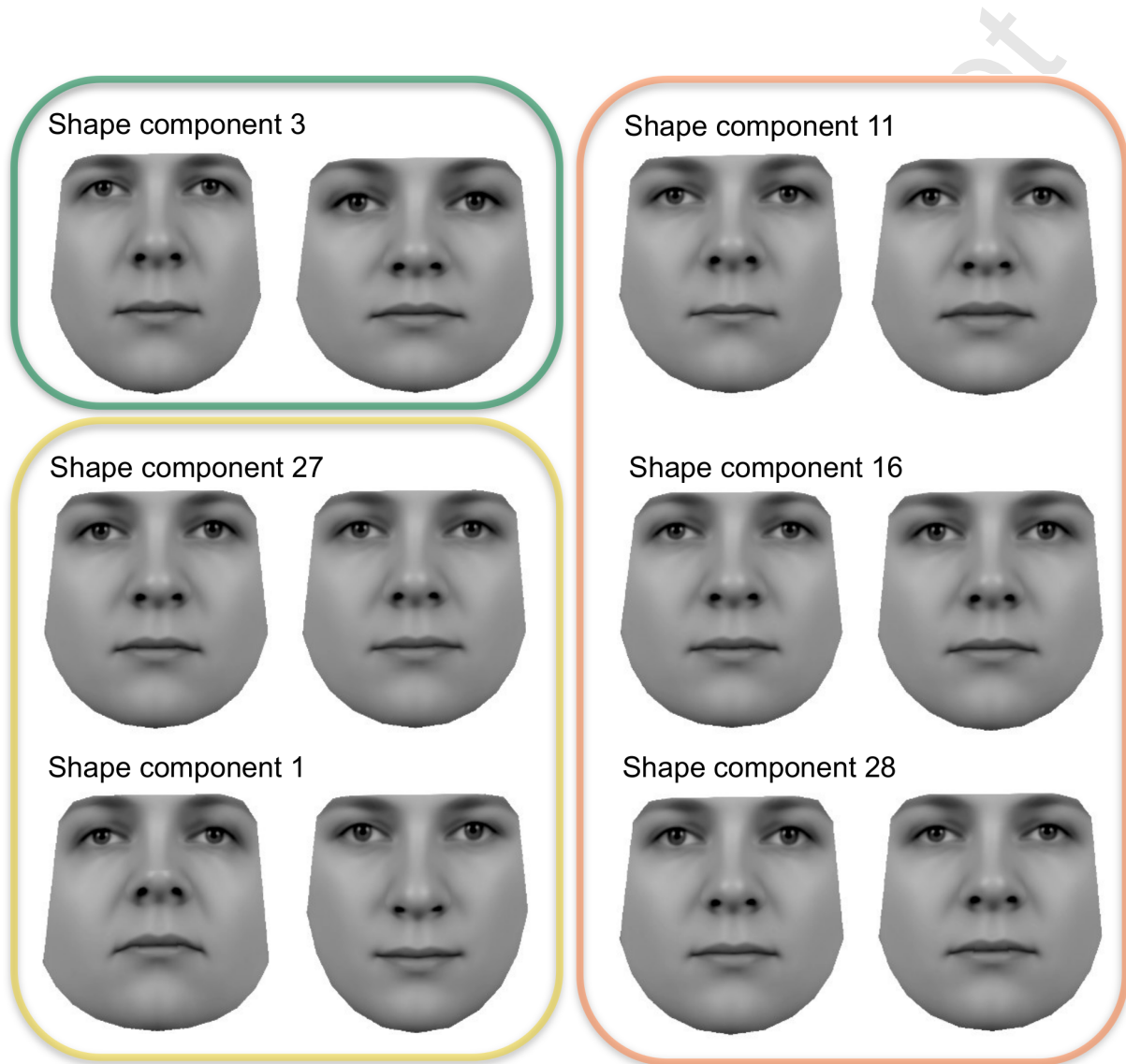


Figure 4. Shape components for the three groups all (green), men (yellow) and women (orange) predicted with statistical significance. The face pairs visualize facial trait components with a distance of ± 2 standard deviations from the mean. Component 3 mainly describes face width; component 27 aspects of mouth shape; component 1 face width (besides head rotation); component 11 mainly size of lips; component 16 distance between eyes and eyebrow width; and component 28 eye size.

distance, eye size and eyebrow width as predictable to a smaller degree. The fact that we achieve a better performance when predicting on genders separately could be caused by differences in facial features between men and women. Generating shape and genetic components for men and women separately instead of controlling for gender effects could lead to a higher prediction accuracy due to a more clearly defined phenotype.

An important fact to consider is the polygenic nature of facial traits, which are likely influenced by thousands of genetic variants. Such a large number of significant SNPs could therefore benefit from being combined to get more clearly defined and reliable predictors for facial features. We attempted this by a PCA, a linear dimension reduction, however, other non-linear dimension reduction techniques could in the future be shown to perhaps better grasp the interplay between SNPs. Also pre-selecting SNPs in genes part of pathways hypothesized to affect cranio-facial development could reveal itself as more appropriate for the achievement of clearer facial trait predictions. Not many studies have at this time attempted to predict facial traits from SNPs making it difficult to compare our results to others. One relevant study predicted from a smaller number of pre-selected SNPs 2.04 % of the variance in nose width in a german cohort and 0.28 % of the variance in bizygomatic distance in a dutch cohort [15]. In comparison we were able to predict 4.4 to 9.6 % of the variance in our shape components by combining a large number of associated SNPs.

This study has several limitations and is only a first step towards facial trait prediction from genome-wide SNP data. Image data was recorded in such a way that we had to disregard texture information and absolute head size information. Additionally, our sample size was small for the prediction of a complex trait; the GWAS' were run on only 827 individuals, which could have resulted in important SNPs not being detected. We then chose to discard a large part of our sample after initial model generation to ensure full separation between feature selection and prediction. Future studies would benefit from using the entire sample throughout the analyses by moving all steps into the cross-validation. Such an approach was not possible in this study because of a then too excessive computational demand. It should also be noted that our results are only applicable to Icelandic individuals, which may have led to inflated predictions due to a larger genetic similarity and thereby larger homogeneity in appearance in the Icelandic population compared to other populations. In the future it would be necessary to apply this method to diverse populations to properly assess generalizability of findings.

For facial trait prediction to succeed, finding the most appropriate dimension reduction and prediction

192 methods, acquiring standardized, preferably 3D, facial images and gathering larger amounts of data is only
193 part of the obstacle. This, because genotype data alone will only enable incomplete prediction of most
194 traits, as seen by approaches combining information from large amounts of common SNPs to gauge the
195 heredity of traits [14,31]. Therefore, epigenetic and environmental data would be of interest to perhaps
196 move towards success in the area of facial trait prediction [32].

197 In conclusion, we have in this study shown that specific facial features can be predicted to a small
198 extent from genetics, even for traits without known candidate genes. It is necessary to move away from
199 approaches using SNPs as single entities and we find our method to hold definite potential for facial trait
200 prediction in forensic science. However, a larger data set and refinement of the used methods is needed
201 before it will be possible to determine to which extent the coding of facial features can be extracted from
202 genome-wide SNP data.

203 **7 Acknowledgments**

204 The research leading to these results has received support from EU funded FP7-People-2011-IAPP grant
205 PsychDPC (GA 286213). THP was supported by The Danish Council for Independent Research Medical
206 Sciences and The Alfred Benzon Foundation.

References

1. Baranzini S, Mudge J, van Velkinburgh J, Khankhanian P, Khrebtukova I, et al. (2010) Genome, epigenome and RNA sequences of monozygotic twins discordant for multiple sclerosis. *Nature* 464: 1351–1356.
2. Sulem P, Gudbjartsson D, Stacey S, Helgason A, Rafnar T, et al. (2007) Genetic determinants of hair, eye and skin pigmentation in Europeans. *Nature Genetics* 39: 1443–1452.
3. Kenny E, Timpson N, Sikora M, Yee M, Moreno-Estrada A, et al. (2012) Melanesian blond hair is caused by an amino acid change in TYRP1. *Science* 336: 554.
4. Eriksson N, Macpherson J, Tung J, Hon L, Naughton B, et al. (2010) Web-based, participant-driven studies yield novel genetic associations for common traits. *PLoS Genetics* 6: e1000993.
5. Medland S, Nyholt D, Painter J, McEvoy B, McRae A, et al. (2009) Common variants in the trichohyalin gene are associated with straight hair in Europeans. *American Journal of Human Genetics* 85: 750–755.
6. Sturm R, Duffy D, Zhao Z, Leite F, Stark M, et al. (2008) A single SNP in an evolutionary conserved region within intron 86 of the *HERC2* gene determines human blue-brown eye color. *American Journal of Human Genetics* 82: 424–431.
7. Eiberg H, Troelsen J, Nielsen M, Mikkelsen A, Mengel-From J, et al. (2008) Blue eye color in humans may be caused by a perfectly associated founder mutation in a regulatory element located within the *HERC2* gene inhibiting *OCA2* expression. *Human Genetics* 123: 177–187.
8. Kayser M, Liu F, Janssens AC, Rivadeneira F, Lao O, et al. (2008) Three genome-wide association studies and a linkage analysis identify *HERC2* as a human iris color gene. *American Journal of Human Genetics* 82: 411–423.
9. Keating B, Bansal A, Walsh S, Millman J, Newman J, et al. (2013) First all-in-one diagnostic tool for DNA intelligence: genome-wide inference of biogeographic ancestry, appearance, relatedness, and sex with the Identitas v1 Forensic Chip. *International Journal of Legal Medicine* 127: 559–572.

10. Andersen J, Johansen P, Harder S, Christoffersen S, Delgado M, et al. (2013) Genetic analyses of the human eye colour using a novel objective method for eye colour classification. *Forensic Science International: Genetics* 7: 508–515.
11. Hemani G, Yang J, Vinkhuyzen A, Powell J, Willemsen G, et al. (2013) Inference of the Genetic Architecture Underlying BMI and Height with the Use of 20,240 Sibling Pairs. *American Journal of Human Genetics* 93: 865–875.
12. Lango Allen H, Estrada K, Lettre G, Berndt S, Weedon M, et al. (2010) Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* 467: 832–838.
13. Manolio T, Collins F, Cox N, Goldstein D, Hindorff L, et al. (2009) Finding the missing heritability of complex diseases. *Nature* 461: 747–753.
14. Yang J, Benyamin B, McEvoy B, Gordon S, Henders A, et al. (2010) Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics* 42: 565–569.
15. Boehringer S, van der Lijn F, Liu F, Gunther M, Sinigerova S, et al. (2011) Genetic determination of human facial morphology: links between cleft-lips and normal variation. *European Journal of Human Genetics* 19: 1192–1197.
16. Paternoster L, Zhurov A, Toma A, Kemp J, St Pourcain B, et al. (2012) Genome-wide association study of three-dimensional facial morphology identifies a variant in PAX3 associated with nasion position. *American Journal of Human Genetics* 90: 478–485.
17. Liu F, van der Lijn F, Schurmann C, Zhu G, Chakravarty M, et al. (2012) A genome-wide association study identifies five loci influencing facial morphology in Europeans. *PLoS Genetics* 8.
18. Claes P, Liberton DK, Daniels K, Rosana KM, Quillen EE, et al. (2014) Modeling 3d facial shape from dna. *PLoS Genetics* 10: e1004224.
19. Stefansson H, Meyer-Lindenberg A, Steinberg S, Magnusdottir B, Morgen K, et al. (2014) CNVs conferring risk of autism or schizophrenia affect cognition in controls. *Nature* 505: 361–366.
20. Holm H, Gudbjartsson DF, Sulem P, Masson G, Helgadóttir HT, et al. (2011) A rare variant in MYH6 is associated with high risk of sick sinus syndrome. *Nature Genetics* 43: 316–320.

21. Marchini J, Howie B, Myers S, McVean G, Donnelly P (2007) A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet* 39: 906–913.
22. Zhang Z (2004) Camera Calibration, Prentice Hall Professional Technical Reference, chapter 2. *Emerging Topics in Computer Vision*. pp. 4–43.
23. MATLAB (2010) R2010b. Natick, Massachusetts: The MathWorks Inc.
24. Viola P, Jones M (2004) Robust Real-Time Face Detection. *International Journal of Computer Vision* 57: 137–154.
25. Cootes T, Edwards G, Taylor C (2001) Active Appearance Models. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 23: 681–685.
26. Fagertun J (2005) Face Recognition. *M.Sc.Eng. Thesis*, Technical University of Denmark.
27. Goodall C (1991) Procrustes methods in the statistical analysis of shape. *Journal of the Royal Statistical Society* 53: 285–339.
28. Bair E, Hastie T, Paul D, Tibshirani R (2006) Prediction by supervised principal components. *Journal of the American Statistical Association* 101: 119–137.
29. R Development Core Team (2012). R: A language and environment for statistical computing. <http://www.R-project.org/>.
30. Burton AM, Bruce V, Dench N (1993) What's the difference between men and women? Evidence from facial measurement. *Perception* 22: 153–176.
31. Vinkhuyzen A, Pedersen N, Yang J, Lee S, Magnusson P, et al. (2012) Common SNPs explain some of the variation in the personality dimensions of neuroticism and extraversion. *Translational Psychiatry* 2: e102.
32. Flintoft L (2013) Enhancers 'fine-tune' face and skull shape. *Nature Reviews Genetics* 14.